Peer-review report of

Tsomokos, D. I. & Flouri, E. (2023). Superior social cognitive abilities in childhood are associated with better reward-seeking strategies in adolescence: evidence for a Social-Motivational Flexibility Model. *advances.in/psychology, 1*(1) https://doi.org/10.56296/aip00002

Round 1

Dear Authors,

Thank you for submitting this interesting manuscript to advances.in/psychology. I have secured the reviews from two experts – one in developmental psychology and another in social cognition. Both the Reviewers and I see several strengths in the manuscript. At the same time, we believe that there are ways to improve it further. I therefore invite you to submit a revised version of the manuscript including a revision letter that outlines the changes.

Both reviewers provided detailed and constructive feedback, and I am grateful for their thorough evaluations. Reviewer 1 asks for the inclusion of more recent literature in the introduction and for more details regarding various aspects of the theoretical frameworks and predictions this study is based on. Reviewer 2 provides excellent suggestions for improving the paper's overall framing. I agree that the paper is best suited to test one model rather than to compare accounts from two models, and I suggest you frame the introduction accordingly.

Both reviewers provide essential suggestions for improving the methods and results sections. For instance, I concur with Reviewer 1's request for more statistical details. In addition to their comments, please provide the following:

- A formal test of whether values are missing at random.
- Please do not present NAs in the violin charts as they are not very informative and distract from the comparison of interest.
- Please provide a fully crossed correlation table (all variables by all variables).
- For Table 3, please present exact p-values and 95% confidence intervals for the regression coefficients. I would also like to see standardized coefficients added to that table so that readers get a sense of the effect size.
- Also, as Reviewer 1 suggests, please present the F tests for the models.
- Please use *B* for unstandardized coefficients and β for unstandardized coefficients.
- In Table 1, within the analytic sample, the SAT scores add up to 99% (likely due to missing values – this is ok). In the "Rest of MCS6" it adds up to 36%. What is the reason for the divergence in these percentages in both groups? Is this a typo?
- Please also provide more information about the weighting procedure.
- Please describe in a bit more detail the test of participants' verbal ability.
- I had difficulty understanding why the disadvantaged stratum did not include wards with more minority-group members. Why were these dimensions

expected to be completely orthogonal? I understand that this cannot be changed now, but readers may want to know.
- Please also indicate the number of the NHS ethics approval.
- Please make the anonymized data and code available to researchers via a repository. Please explain in the case this is not possible.

I also concur with Reviewer 2, who asks for a more fine-grained analysis of the data and justification of it, both when it comes to the IV and DV. You are presenting very unique data, but many details currently remain hidden. I leave it up to you how you address the issues highlighted by Reviewer 2, but I ask for a more detailed analysis that can complement the current results and a sound justification for how you coded the IVs and DVs.

Both reviewers also provide important ways to improve the discussion. Both ask you to engage further in the strengths and limitations of the paper. Reviewer 1 asks for a summary of the results at the end of the manuscript. Reviewer 2 asks for a section that outlines in more detail the theoretical and practical implications of your work.

To sum up, we all believe that this paper has the potential to make a strong contribution to the field pending a set of revisions. When submitting your revision, please attach a revision letter (in PDF format), that provides an itemized overview of how you addressed each of the points highlighted by the reviewers and the additional points outlined in this letter. Please submit the revision within 45 days (within January 15).

Best,
Jonas R. Kunst
Editor-in-Chief

Reviewer 1:

In this paper, the authors examine the association between the onset of Theory of Mind (ToM) abilities at ages 5 and 7 using the Sally-Anne task, and risk adjustment at age 14, measured using the Cambridge Gambling Task (CGT). After controlling for possible confounders, the authors find that children who exhibited ToM abilities in the Sally-Anne task at ages 5 and 7 were better able to adjust their betting behaviour to account for the likelihood of being rewarded on the CGT. The authors suggest that ‘superior’ social cognition in childhood supports better risk adjustment in adolescence.

There are several strengths to this paper, including the large sample size and inclusion of possible confounders in the analyses. In addition, the perspective is relatively novel, as there is limited work examining how social abilities in childhood relate to decision-making in adolescence. However, I think there are areas where the paper could be improved, particularly in the framing of the study in the Introduction, and the contextualisation of the findings in the Discussion. By revising these sections to provide stronger theoretical framing, I believe the contribution of this paper to the literature will be clearer. I have outlined specific suggestions below.

Introduction:

The review of the adolescent risk taking literature could include more recent papers to better reflect the current state of the field. Two recent theory papers come to mind in particular, as they outline how social cognition during adolescence may relate positive risk taking behaviours, which is closely related to the topic of this paper. Integrating these papers into the literature review would further strengthen the rationale for the current study. The references for these papers are:

Do, K.T., Sharp, P.B., & Telzer, E.H. (2020). Modernizing conceptions of valuation and cognitive control deployment in adolescent risk taking. *Current Directions in Psychological Science, 29*(1), 102-109.

Duell, N., & Steinberg, L. (2021). Adolescents take positive risks, too. *Developmental Review*, *62*, 100984.

Page 3, line 52: There are quite good estimates for the period during which ToM abilities emerge in neurotypical children. Therefore, I would suggest the authors are more specific about these ages rather than stating these abilities emerge “At some point”.

When discussing the procedure by which the Sally-Anne task (SAT) was administered in the MCS, it would be helpful for the authors to identify the specific ability or abilities this demanding version of the task relies on (in addition to ToM) to justify that it is superior social cognition that is in fact measured by this task. Earlier in the introduction, it is stated that the SAT requires “executive functions such as response-selection, response-inhibition, and working memory” which could suggest that the children who correctly responded to the demanding version of the SAT are those with greater executive functioning skills. As executive functioning predicts greater risk adjustment (e.g., Bell et al., 2019), the argument could be made that the children who respond correctly to the version of the SAT employed in the MCS have improved executive functioning skills, and this drives the association between their performance on the SAT and risk adjustment in adolescence. Clearly defining the social skills measured in the demanding version of the SAT will help the authors to exclude this alternative explanation.

The rationale for why childhood social cognition would predict adolescent risk taking could be strengthened. Currently the two sections on adolescent risk taking and ToM are quite distinct, but an additional paragraph that integrates these two topics and explains how social cognition is important for adjusting risk behaviour would help to frame the study better.

One aspect of the argument that is implied, but should be stated explicitly, is that individual differences in ToM in childhood predict social cognition in adolescence. For example, the authors could cite evidence that the timing with which children acquire ToM predicts social skills during adolescence.

The authors should have a clear statement of their hypothesis in addition to the research question.

Throughout the introduction, it would be helpful to mention that risk taking and ToM are associated with some of the measures controlled for in this study, as there is limited justification for the specific confounders in the current version of the manuscript. For example, there is some evidence that socioeconomic status is associated with risk taking in adolescence (Brieant et al., 2021). The rationale for these confounders only needs to be brief but would further justify the study design.

Methods:

In their description of the CGT, it would be helpful to be more explicit about how the RA measure is converted to a numerical value. Does this value refer to the number of points the participant retains at the end of the study?

Given my earlier point about other possible explanations for performance on the SAT predicting later risk adjustment (i.e., that the SAT is measuring executive functions), the authors should consider entering the CANTAB spatial working memory as an additional control variable in their regression model. This measure was also measured during the sweep where the CGT data were collected https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-Guide-to-the-Datasets-022014.pdf

Can the authors state whether any participants were excluded from the final sample? For example, did they exclude participants with a diagnosis of Autism Spectrum Disorder?

Results:

The results are generally reported clearly. However, it would be important to include the F statistic for each of the models in addition to the R2 value, as well as the 95% confidence intervals for their regression models.

Further, it should be noted in the text that the overall models were not significant, meaning that together the predictors did not explain a significant proportion of variance of participants' risk adjustment. This is an important caveat to the findings, although the individual predictors were significant.

Discussion:

Page 16 line 286-287: it would be helpful to define the specific social skills measured by this demanding version of the SAT.

Throughout, the discussion could be improved by referring to more recent literature on social cognition and risk taking in adolescence, such as the Do et al. (2020) and Duell and Steinberg (2021) papers referenced above.

One major limitation that is not currently acknowledged is that the CGT does not contain any social stimuli. In lines 300-303, the authors purpose that adolescent neural development leads to an improved ability to adjust one's goals and motivation to adapt to their changing social environment. However, the behavioural task they use to test this hypothesis is an asocial risk taking task and therefore there is a disconnect between the task used in this study and the purported explanation for this behaviour. While I appreciate the authors are restricted by the methods used in the MCS, it would be important to recognise the limitations of this task to test the authors' hypothesis and identify this as an avenue for future research.

The discussion ends quite abruptly – a paragraph summarising the findings would help to conclude the study.

Minor: "mature brain system" could be phrased more concisely to refer to the specific developments that occur during adolescence (i.e., the rapid development of the dopaminergic reward system and protracted development of regions responsible for cognitive control).

References:

Bell, M. D., Imal, A. E., Pittman, B., Jin, G., & Wexler, B. E. (2019). The development of adaptive risk taking and the role of executive functions in a large sample of school-age boys and girls. *Trends in neuroscience and education*, *17*, 100120.

Brieant, A., Peviani, K. M., Lee, J. E., King-Casas, B., & Kim-Spoon, J. (2021). Socioeconomic risk for adolescent cognitive control and emerging risk-taking behaviors. *Journal of research on adolescence*, *31*(1), 71-84.

Reviewer 2:

Review of "Superior social cognitive abilities in childhood are associated with better reward-seeking strategies in adolescence: evidence for a Social-Motivational Flexibility Model"

I read this manuscript with great interest. The issue it addresses – how social cognitive scores at ages 5 and 7 predict risk adjustment at age 14 – is stimulating, with important theoretical as well as practical implications. The work is based on a unique, large, appropriate database, which is skillfully analyzed. The results are stimulating, and both the conclusions and limitations are laid out in a fair and balanced way. Overall, I believe that that this manuscript may be a strong candidate for publication in *Advances in Psychology*, pending appropriate revision.

My recommendations for revision involve aspects of the framing, questions about the data, and few suggestions about limitations.

With respect to the framing, the implication from the way the Introduction is organized and the conclusion that is drawn early in the Discussion makes it seem like this work is a critical test of the Dual Systems Model versus Social-Motivational Flexibility Model of adolescent development. However, it seemed to me that the data analyzed was appropriate for testing a key hypothesis of the Social-Motivational Flexibility, but the data do not directly speak in a comprehensive way to the core constructs outlined in the Dual Systems Model. Thus, I was not convinced that the current work represents a true critical test, and that concerned me about the meaning and contribution of the current research. Thus, I suggest some revision in the framing. What I think the current work does well is test a key hypothesis derived from the Social-Motivational Flexibility Model. So, my recommendation is to begin with a description and emphasis of that model. (The first two paragraphs of the Introduction currently focus on the Dual Systems Model.) After describing the Social-Motivational Flexibility Model in a way that provides a conceptual context for the main hypothesis being empirically test, a paragraph could be included about how the Social-Motivational Flexibility Model contrasts with an alternative prominent model, the Dual Systems Model, generally and perhaps identify how and why the Dual Systems model would suggest a different result than the Social-Motivational Flexibility Model in this specific context. This reorganization of the Introduction would foreground the Social-Motivational Flexibility Model, which I believe is the theoretical backbone of the current study, and relieve the researchers from having to add more variables to the analyses to conduct a more comprehensive critical test of the two models. My view of the current work is that while it makes a strong case for the Social-Motivational Flexibility Model through empirical support for a hypothesis directly derived from it, it does not necessarily provide a full test of the alternative, Dual Systems Model. Thus, my recommendation is to emphasize more what the work does best – which seems like the main objective given the design of the study – and make consideration of Dual Systems Model secondary in the Introduction. The authors might, however, amplify some of their consideration in the Discussion of why the Social-Motivational Flexibility Model provides a better account for the findings than the Dual Systems Model, perhaps important some material that is currently in the Introduction.

Although I found the data reported to be relevant to the main hypothesis and well-analyzed, I had some questions and suggestions for further analyses, which I think could produce an even more persuasive case for the authors' conclusions. One point relates to the outcome variable; the other to a predictor variable. In terms of the outcome variable, I agree that the Risk Adjustment score from the Cambridge Gambling Task is most relevant to testing the hypothesis derived from the Social-Motivational Flexibility Model. However, my understanding is that there are other measures that can be derived from that task, including for example quality of decision making, risk taking, delay aversion. Some of these seem relevant to expectations related to the Dual Systems Model. So, if the authors continue to make the effectiveness of the Social-Motivational Flexibility Model relative to the Dual Systems Model a point of emphasis in the manuscript, even if mainly in the Discussion, then I recommend including additional, relevant outcomes measures particularly relevant to the comparison of the models. Also, are any of these measures especially relevant to testing the Social-Motivational Flexibility Model (perhaps even as a foil to the risk adjustment measure, to show, for example, that social-cognitive functioning at an earlier age distinctively predicts risk adjustment, for

example in ways over and above or in contrast to overall risk taking? Providing an appropriate, broader pattern of data, rather than a single outcome measure, could strengthen the authors' case with even more compelling evidence.

Regarding the predictors, the authors basically contrast those who "pass" the Sally-Anne Task at both ages 5 and 7 to those who do not. Is there any precedent for such categorization, this is having passed it for both versus having achieved a passing performance by age 7, the time most proximate to the measure of risk adjustment? If so, that would be valuable information to include in support of the current analytical strategy. Alternatively, or perhaps in supplementary analyses, the authors might examine all four Sally-Anne Test outcomes – (1) fail at both ages 5 and 7, (2) fail at 5 but pass at 7, (3) pass at 5 but fail at 7, and (4) pass at both 5 and 7 – and consider those results. For example, if failing at age 5 but passing at age 7 produces similar results as passing at both, that could be very informative developmentally. It seems that the authors are leaving valuable information "on the table."

Turning to the Discussion, addressing the points I have raised would likely produce changes in the Discussion, for further evaluation of a revised manuscript. Independent of those changes, there is a potential issue that I feel merits further consideration as a limitation. At the beginning of the Results section, the authors commendably discuss Sample Bias. The authors report that the sample analyzed differed in a number of ways "compared to the rest of the cohort at age 14." On the one hand, it might seem that the sample differs systematically enough from the rest of the cohort that the generalizability of the findings might be questioned. I recommend raising this as another potential limitation. However, the authors might then note that the MCS sampling frame "was designed to over-represent families living in certain wards, namely, (a) wards in the UK with high child poverty, (b) wards in England with higher concentration of ethnic minorities, and (c) the smaller UK countries 119 (Plewis et al., 2004)." If possible, the authors might also, if data were available, compare the sample they analyzed to other norms for the ages studied for representative samples. I also recommend the authors add a paragraph in the Discussion about the practical implications of the work, perhaps with respect to parenting, socialization, and/or education (including suggestions about possible interventions).

So, I believe that this manuscript represents the core of important and skillfully-conducted research that has the potential to merit publication in *Advances in Psychology*. The comments and suggestions that I offer are intended to make a promising manuscript even stronger and more persuasive. I think the case for support for the Social-Motivational Flexibility Model is sufficient, even without making a case from a critical test compared to the Dual Systems model, and I urge the authors to make that the focus of the current work. I believe that the authors can address the issues I have raised with some re-thinking, re-framing, and re-writing, along with the addition of some data (if available) and some re-analysis. I encourage the authors to revise the manuscript and resubmit it for further consideration for publication in Advances in Psychology. The work is stimulating, and the results are potentially both theoretically and practically important.

Round 2

Dear Authors,

Thank you for submitting a revised version of your manuscript for consideration at advances.in/psychology. I was fortunate to obtain reviews from the same two reviewers who read the first version of the manuscript. I also read the manuscript independently of their reviews. We all applaud you for addressing most concerns in the letter and in the manuscript. Both reviewers suggest a set of minor revisions that would prepare this piece for publication. Please address each of their points in a revised submission.

You responded well and in detail to most of our comments. However, I concur with Reviewer 2 that much of the information provided in the letter would be helpful in the main manuscript. Thus, I ask you to move the information they request to the main text. Please also move the following information to a supplementary file uploaded at OSF and refer to it in the main manuscript:

1. Missing value analyses

2. Analyses with standardized variables. In addition to moving the results from these tests to the SOM, I found the following text striking, which I suggest you add to the main manuscript in the following way: “Analyses with standardized variables showed that the relative impact of our independent variable, SAT (allQ: Pass), is roughly 35% larger than that of income (FOECDUK0) and maternal education (FDACAQ00), and 25% larger than that of mental health, as measured via total SDQ scores (FEBDTOT; see SOM for details).”

Further, I understand that it space-wise was hard to present exact p-values in Table 3. In the text, you present the exact p-value for the most critical test of interest (the effect of “pass”) for Model 3. Please also present in the text p-values and 95% CIs for this effect for Models 2 and 3.

Finally, I agree with Reviewer 2 that you should provide results from tests with the other CGT outcomes in the main text (i.e., the statistically non-significant findings that you currently briefly summarize in the letter). It is then essential to distinguish these outcomes already in the introduction and explain which was considered particularly relevant to the research question. It seems as if you interpret the non-significant findings for the other outcomes as a type of discriminant validity, but it would be beneficial for a reader if you discussed this explicitly.

Small last issue: Please provide a page number for the quote in lines 335-338.

I look forward to receiving your revisions within 30 days. I intend to make a decision on the next version of the manuscript without further consulting with the two Reviewers.

Best,

Advances.in
Reinventing Academic Publishing

Jonas R. Kunst

Editor-in-chief

Reviewer 1:

I appreciate the authors' thorough responses to the issues raised and these points have been adequately addressed in the revised manuscript. The theoretical framing of the study is much stronger and as such, the contribution of this study to the literature is clear. In addition, the description of the methods and results have been much improved in the revised manuscript. I have some minor suggestions based on the revised manuscript that I would encourage the authors to consider:

Introduction

Page 3 lines 65-67: It is unclear why it is important that the vmPFC has been implicated in false beliefs, irony and lying. It seems more pertinent that neural regions implicated in adjusting attention overlap with ToM abilities (a point raised in line 62). Perhaps remove this line or explain in greater detail why it is a crucial piece of information?

Methods

The description of the CGT and how the outcome measure was derived is much clearer. Could the authors report the number of trials in the CGT in addition to the information they currently include?

Page 8 line 194: I do not think it is correct to describe the operationalisation of maternal education as an interval variable. Interval variables are numerically ranked and have mathematically equal spacing between points on the scale. I do not think this description applies to the operationalisation of material education, as the difference between no education (i.e., a value of 1) and level 1 of national qualifications (e.g., GCSEs grades 3, 2, 1 or D, E F, G) is not mathematically equal to the difference between level 1 and level 2 qualifications (e.g., GCSE grades 9, 8, 7, 6, 5, 4 or A*, A, B C). I believe this variable would be more appropriately described as an ordinal variable.

Page 10, line 241: The citation for R should read "All calculations were performed using R (R Core Team, 2020)…". I also suggest the authors report the version of R used when the analyses were conducted, as there have been updates to the software. The reference for R should also be corrected to 'R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/'.

Results

The results have been reported clearly and contain all the appropriate information.

Discussion

While I appreciate the authors' revisions to the Discussion, I still think there is scope to further contextualise the findings in relation to recent empirical research on social cognition in adolescence and how this relates to risk taking/adjustment. For example, Reiter and colleagues (2019) have demonstrated that the degree to which adolescents adjust their risk taking behaviour to reflect that of their peers is associated with their real-life social integration. These empirical findings complement those reported in the current study and highlighting the emerging evidence in this area would further strengthen the Discussion.

Reference:

Reiter, A. M. F., Suzuki, S., O'Doherty, J. P., Li, S.-C., & Eppinger, B. (2019). Risk contagion by peers affects learning and decision-making in adolescents. *Journal of Experimental Psychology: General, 148*(9), 1494–1504. https://doi.org/10.1037/xge0000512

Reviewer 2:

I was Reviewer 2 for the original submission of this manuscript. I reviewed this revised version of the manuscript both before and after reading the authors' Response Letter. I believe that the authors have considered the feedback received from me, Reviewer 1, and the Editor, and that the revised text addresses many of the core issues. For instance, consistent with my recommendation, the authors now focus the development of the rationale for the current research on the Social-Motivational Flexibility Model rather than on a comparative analysis of that model compared to the Dual Systems Model. This new framing is clear and effective.

However, there were a few issues that I originally raised that the authors spoke to in their Response Letter but were not addressed in the manuscript itself. I think that other readers may have similar questions, so I am recommending that the authors import some material from the Response Letter directly into the text of the manuscript. I will defer to the Editor, though, if the Editor believes that incorporating that information is more distracting than helpful.

One of these points involves the authors' decision to consider as a "pass" on the Sally-Anne Task a person who passed at both at ages 5 and 7. I agree with the authors that it would "be practically impossible to decipher in a operational manner" some mixed pass-fail results. I do, however, think that a fail at 5 and pass at 7 could be rather straightforwardly attributed to developmental factors. A pass at 5 and fail at 7 would be hard to explain developmentally but could speak to issues of reliability of the measure. Rather than go back and forth on this issue, though,my recommendation would be simply to have the authors explicitly describe the basis of their dichotomization decision and, if possible, cite a research precedent supporting that. Currently, the authors state (p. 8, lines 171-173), without providing a rationale: "We consider a child to have fully passed the test if Q1 and Q2 were answered correctly during both sweeps." While the authors' dichotomization decision is reasonable, I believe that readers would benefit from seeing the reasoning behind

it. This is a pivotal decision that determines the analyses, results, and interpretation of the data.

Another point involves the issue I raised about sample bias. I found the explanation in the Response Letter to be very helpful. The authors wrote, “The sample bias observed in our study is typical of analytic samples in MCS that are selected on the basis of there being valid responses in cognitive tasks, such as the Cambridge Gambling Task used in our work. The primary reasons for this are attrition over time, and a preference of certain families to participate in cognitive assessments…. As a result of this sampling in MCS, our eventual analytic sample is representative of the general UK youth population, although slightly more representative of families with higher socioeconomic status.” I recommend that the authors add a few sentences to the manuscript about this issue.

While I think the first two issues can be addressed by simply adding a few sentences here and there, the third and last issue will take more work. When the original manuscript focused on comparing the Social-Motivational Flexibility Model with the Dual Systems Model, I asked about other relevant measures besides risk adjustment. As the authors explain in their Response Letter, “We have shifted the main focus of the manuscript to the Social-Motivational Flexibility Model as such, placing less emphasis on its comparison against the Dual Systems Model, as explained in our response to the previous item. In both cases, however, the other measures of the Cambridge Gambling Task are slightly outside the scope of the present study, as they pertain to constructs such as delay aversion and deliberation time, which we have now checked and observed that SAT does not predict.” I am satisfied with this explanation. However, it might be useful to add such a statement about the other measures and their results in Supplementary Materials.

The authors also noted that the overall risk-taking measure is potentially relevant, and in the Response Letter they present the results of the analysis showing that the Pass-Fail does not significantly predict overall risk taking. I think the finding that Pass-Fail predicts risk adjustment but not overall risk-taking is important information for readers to see in the main text. It shapes the take-home message in a significant way. Including this information in the main text, which I strongly recommend, may also require some refinement of the rationale presented. For instance, on page 2 of the manuscript, the authors write, “Therefore, the Dual Systems Model provides a compellingly simple account for the fact that, in general, risk-taking increases from late childhood to middle adolescence, and then gradually decreases again, while the self-regulatory control system continues to mature into adulthood (Shulman et al., 2016). In addition to increased risk-taking and reduced inhibition, adolescence is also a 30 period of increased social-affective and social-cognitive maturation.” It would thus be useful to more explicitly distinguish overall risk-taking from risk adjustment conceptually early on, and then to develop the rationale from the Social-Motivational Flexibility Model in a way that not only focuses readers on the expectation for risk adjustment (as the manuscript currently does) but also as distinct from overall risk-taking. This distinction would then also merit some attention in the Discussion section. I believe that this would be a very valuable addition to the current version of the manuscript.

In conclusion, I commend the authors for their responsiveness to the feedback they received in the initial round of review. I believe that the changes they have made have substantially strengthened the work. I am even more. Convinced than before that this work will eventually merit publication. However, as I have outlined, a few additional changes would make a strong manuscript even stronger.